Step-by-Step Guide: Integrating Looker Studio with a Project Website using ETL Pipeline

Overview:

This guide walks through how to integrate a real-time data dashboard (Looker Studio) into a website using an ETL pipeline that moves data from MongoDB Atlas → Python → BigQuery → Looker Studio → Website Embed.

Step 1: Prepare Your MongoDB Data

1. Connect to MongoDB Atlas

- - Set up a MongoDB Atlas cluster.
  - Get your connection string (e.g. mongodb+srv://<user>:<pass>@cluster.mongodb.net/myDB)

2. Store Weekly Data

- - Upload weekly scraped Coles/Woolworths data using Python (e.g., using pymongo).
  - Ensure consistent schema (e.g., product name, price, brand, category, store, date).

Step 2: Extract & Transform Data with Python

1. Create ETL Script

- - Use pymongo to extract data from MongoDB.
  - Clean the data using your refined synthetic data cleaning tool (nulls, outliers, schema validation).
  - Convert the cleaned data to a Pandas DataFrame.

2. Load into BigQuery

- - Use google-cloud-bigquery library to upload your DataFrame to a BigQuery table.
  - Make sure table schema matches Looker Studio needs.

Step 3: Connect BigQuery to Looker Studio

1. Create a New Dashboard

- - Open Looker Studio.
  - Select BigQuery as your data source.
  - Choose your dataset and table.
  - Add charts, filters, tables as needed for the visualisation.

2. Customize the Dashboard

- - Add branding, filters (e.g., by store or product), and date ranges.
  - Test with sample data to ensure real-time sync.

Step 4: Embed Looker Studio Dashboard into Website

1. Get Embed Code

- - In Looker Studio: File → Embed report → Enable embedding.
  - Copy the iframe HTML snippet provided.

2. Embed into HTML

Use this in your project's HTML file frontend:
```
<iframe width="100%" height="600"
 src="https://lookerstudio.google.com/embed/reporting/your-report-id/page/xyz"
 frameborder="0" style="border:0" allowfullscreen></iframe>
```

Step 5: Automate and Maintain

1. Schedule ETL Runs

- - Use cron, Airflow, or Google Cloud Scheduler to run your ETL script weekly.

2. Update Dashboard

- - Looker Studio auto-updates when BigQuery data changes—no need to refresh manually.